Price impact versus bid-ask spreads in the index option market

Article (Published Version)


This version is available from Sussex Research Online: http://sro.sussex.ac.uk/id/eprint/101777/

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher’s version. Please see the URL above for details on accessing the published version.

Copyright and reuse:
Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.
Price impact versus bid–ask spreads in the index option market

Andreas Kaeck a, Vincent van Kervel b, Norman J. Seeger c,∗

a University of Sussex Business School, Sussex House, Falmer, Brighton, BN1 9RH, United Kingdom
b Escuela de Administración, Pontificia Universidad Católica de Chile, Avenida Vicuña Mackenna 4860, Macul 7820436, Santiago, Chile
c VU Amsterdam, De Boelelaan 1105, 1081 HV, Amsterdam, The Netherlands

A R T I C L E  I N F O

JEL classification:
G12
G13
G14

Keywords:
Options
Liquidity
Price impact
Informed trading

A B S T R A C T

We investigate the puzzle of why bid–ask spreads of options are so large by focussing on the price impact component of the spread. We propose a structural vector autoregressive model for trades in the option market to analyze whether they move the underlying price and/or the underlying’s volatility. Our model captures cross-option strategies by pooling order flows across contracts after a decomposition into exposure to the underlying asset and its volatility. While our estimates confirm that S&P500 option trades indeed significantly move the underlying and the volatility, the economic magnitudes are very small. Hence, large bid–ask spreads of options remain a puzzle.

1. Introduction

An important puzzle in the option market literature is why the trading costs are so large, as represented by bid–ask spreads. These spreads are large both relative to the value of the option, as well as relative to the liquidity of the underlying. For example, in our sample of SPX options written on the S&P 500 index in 2014, the average effective option spread is $0.59 or 1.69% of its value (Table 1). Muravyev and Pearson (2020) study this puzzle and find that equity option effective spreads are 2.2% relative to option value, which holds after a clever adjustment to the midpoint to reflect that trades are more likely to happen on the ask (bid) when the unobserved fundamental value is higher (lower) than the quoted midpoint.

In this paper, we further investigate this puzzle with a thorough analysis of price impact, which is an important component of the bid–ask spread. We propose a novel methodology to estimate price impact that addresses: i) that options have price impact on both the underlying asset and its volatility; and ii) that trading is spread across a large cross-section of sometimes hundreds of options, which are all traded simultaneously, but differ in strike price, expiration date, and option type (put or call).

Our main finding is that the price impacts in the SPX option market are surprisingly small. We pool the delta component of option trades in the whole cross-section at the hourly level to obtain a net dollar exposure comparable to trading the underlying asset directly, and find that a trade shock of $823 million (one standard deviation in the aggregate SPX option market) increases the underlying by only 5.1 bps.1 This price impact is truly small relative to the massive order flow shock generating $823 million exposure to the S&P 500. While small, the impact is nevertheless fairly precisely estimated with a standard error of 1.15 bps. The

We thank Dmitriy Muravyev, Ioanid Rosu, and Bart Zhou Yueshen for extremely useful comments. We are also thankful for the suggestions by seminar participants at the 2018 Lancaster financial econometrics conference. van Kervel gratefully acknowledges the financial support of the Fondecyt Iniciación, Chile (project 11150485).

* Corresponding author.
E-mail addresses: ak486@sussex.ac.uk (A. Kaeck), vincentvankervel@gmail.com (V. van Kervel), n.j.seeger@vu.nl (N.J. Seeger).

1 We proxy the underlying S&P 500 with the SPY ETF, and find an estimated dollar impact of $0.98, relative to the sample average value of the SPY of $1,930.

https://doi.org/10.1016/j.finmar.2021.100675

Received 6 May 2021; Received in revised form 14 September 2021; Accepted 15 September 2021
Available online 24 September 2021

© 2021 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Identifying whether price impact originates in the underlying asset or the volatility process is challenging due to the strong negative contemporaneous correlation between the two processes [known as the leverage effect (Black, 1976)]. This correlation is \(-0.858\) in our sample at the hourly frequency, and we tackle the issue by re-estimating our main VAR model after filtering out order processing costs, which may be smaller when trades execute at prices inside the quoted bid and ask. We show statistics for the full sample, as well as for subsamples of in-the-money (\(|\Delta| \geq 0.65\)), at-the-money (\(0.65 > |\Delta| > 0.35\)), and out-of-the-money (\(0.35 > |\Delta|\)) with \(|\Delta|\) denoting the absolute value of the Black–Scholes option delta.

## Table 1

Bid–ask spreads of SPX options.

<table>
<thead>
<tr>
<th>Variable</th>
<th>All options</th>
<th>In-the-money</th>
<th>At-the-money</th>
<th>Out-of-the-money</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean</td>
<td>sdev</td>
<td>mean</td>
<td>sdev</td>
</tr>
<tr>
<td>Effective quoted half spread (%)</td>
<td>3.27</td>
<td>3.55</td>
<td>1.29</td>
<td>1.90</td>
</tr>
<tr>
<td>Effective spread (%)</td>
<td>1.69</td>
<td>2.78</td>
<td>0.61</td>
<td>1.35</td>
</tr>
<tr>
<td>Effective quoted half spread ($)</td>
<td>1.21</td>
<td>0.85</td>
<td>1.49</td>
<td>0.69</td>
</tr>
<tr>
<td>Effective spread ($)</td>
<td>0.59</td>
<td>0.67</td>
<td>0.72</td>
<td>0.76</td>
</tr>
</tbody>
</table>

This table shows dollar-weighted means and standard deviations of spreads in the SPX option market over 2014. The effective quoted half spread is defined as the bid–ask spread prevailing just before a trade. The effective spread equals the trade price minus midpoint (signified by the direction of the trade), which may be smaller when trades execute at prices inside the quoted bid and ask. We show statistics for the full sample, as well as for subsamples of in-the-money (\(|\Delta| \geq 0.65\)), at-the-money (\(0.65 > |\Delta| > 0.35\)), and out-of-the-money (\(0.35 > |\Delta|\)) with \(|\Delta|\) denoting the absolute value of the Black–Scholes option delta.

Our results seem to make the option bid–ask spread puzzle even more puzzling. Indeed, it is unlikely that order processing costs explain the puzzle in the current electronic era. Also, the SPX option market is highly competitive so oligopolistic market maker rents are an unlikely explanation. Further, explanations based on option hedging costs are also unsatisfying as the underlying can be hedged by the SPY ETF, which is the world’s most liquid asset.

The main contribution of this paper is a methodological framework that analyzes the price impacts of the order flows of all option contracts jointly. The model first recognizes that any option trade provides exposure, to the underlying asset (through the option delta) and to its volatility (through the option vega). Accordingly, we disentangle the effect of an option trade on the underlying’s price or volatility by constructing two order flow exposures which multiply option net order flow (defined as buyer-originated minus seller-originated volume) by the option delta or vega, which we coin “delta order flow” and “vega order flow”. These order flows can then be meaningfully aggregated across options with different characteristics. We next relate the two option order flows to changes in the underlying price and its volatility in a vector autoregressive (VAR) model. We further add a fifth equation with the net order flow in the underlying asset, which allows for a direct comparison of price impact in the underlying asset and the aggregate option market. We address the challenge that volatility is unobservable by linking a structural option pricing model to the model-free VIX framework. This allows us to link the unobservable volatility to an observable volatility index, which yields consistent delta and vega estimates and neatly integrates into the VAR system. Thus, our framework extends the seminal work of Hasbrouck (1991), who proposes a two-equation VAR model relating order flows in a stock to its price changes.

Our framework yields several advantages over previous studies that typically examine price impact in a single option contract on either the underlying or the volatility. First, by pooling order flow exposures across option contracts, our results represent the magnitude of price impact in the aggregate market. Our results indicate that aggregate price impacts are economically small. In contrast, previous studies typically document the presence of price impact (and therefore informed trading) in a single contract, but the magnitude is not easily translated into an aggregate effect. We effectively apply a data reduction approach by imposing the economic structure from a theoretical option pricing model to summarize the information content in a large number of order flows in just two variables: delta and vega order flow. Second, cross-option strategies are extremely common and may account for more than 75% of volume (Fahlenbrach and Sandás, 2010). We show that the portfolio price impact may be severely reduced when incorporating cross-option price impact, as compared to a naïve approach, which sums the price impacts of the individual option trades. For example, when considering straddle and strangle strategies, we predict a twenty-fold reduction in overall price impact. This reduction is due to the fact that these strategies are close to being delta neutral, and, therefore, the price impact on the underlying largely cancels out. And third, theoretical models with strategic informed traders show that they typically trade in several correlated assets simultaneously (Blais and Hillion, 1994; Boulakov et al., 2012). These theoretical models further motivate our joint analysis of option order flows to measure price impact.

See, for example, Glosten and Harris (1988), Madhavan et al. (1997), and Huang and Stoll (1997).

unexpected, this prediction only goes through the leverage effect. That is, delta and SPY flow increase the underlying and therefore indirectly decrease the volatility. Vega flow, however, still predicts variation in volatility after filtering out the variation explained contemporaneously by the underlying.

Our model imposes a strong economic structure on the data through three crucial assumptions, which we validate with several tests. First, using a Wald test, we examine whether the delta and vega price impacts vary across options with different expiration dates or strike prices. We do not reject equality for vega flows, meaning the vega component of order flows can indeed be pooled across options. The price impact of delta flow does differ significantly across groups of options, and we find it to be higher for options with a short time to maturity and low moneyness. Second, we test the information loss of our model, which uses a two-factor option pricing model (underlying and volatility) to summarize the price changes across all options with different strike prices and maturities. The analysis, based on Bakshi et al. (2000), indicates that our model explains 99.2% of the variation in individual option price changes. Together, the two tests seem to validate our imposed structure, and provide statistical support for a five-equation system to capture the information content of hundreds of option prices and order flows. And third, we test the assumption of the model stemming from the ordering of the equations in the VAR model. In the robustness section, we consider several different orderings of the equations and confirm that all results hold in alternative model specifications.

We are certainly not the first to study the impact of option order flows on the direction of the underlying or volatility (see Footnote ). In particular, Bollen and Whaley (2004) run similar regressions of changes in volatility on net buying pressure of ATM call and put options. Their results, and interpretation, are consistent with inventory models that net order flow has a temporary impact on option implied volatility. This mechanism, together with institutional buying pressure in put index options, is one of the explanations for the volatility smile for example. Gärleanu et al. (2009) formalize this point in a model with risk-averse option market makers who charge (cross-option) price pressures, which in equilibrium are proportional to the unhedgeable component of the option inventory position. These authors do not allow option order flows to move the level of the underlying or volatility, however, which is conceptually important in models of informed trading. Nevertheless, our results do not rule out predictions based on inventory models, because transitory price pressures may exist in the underlying and volatility as well. Rourke (2014) uses a similar price impact VAR model with delta and vega flows estimated using order flows of a single option contract. He uses the returns of a straddle, a portfolio with only vega exposure, to proxy for VIX returns, which makes it difficult to gauge the economic magnitude of the impact of volatility. We contribute to this line of investigation by analyzing order flows in the cross-section of option contracts to address cross-option strategies. We find that cross-option price impact is of first order importance and may explain why the aggregate price impacts turn out to be very small.

This paper proceeds as follows. Section 2 introduces how we extend the Hasbrouck (2003) price impact model to take option markets into account. In Section 3 we describe the empirical setup of the analysis and Section 4 presents the main results of the study. Sections 5, 6, and 7 give results for higher data frequencies, robustness results, and conclusion, respectively.

2. A price impact model for option markets

In this section, we presents a novel methodology to measure the price impact of option order flows. We extend the VAR model of Hasbrouck (1991) to option markets and explicitly take into account (i) the cross-option correlations in order flows, (ii) the cross-option price impacts, and (iii) the fact that options can be used to speculate on both the underlying and its volatility.

Consider an asset on which $N$ different option contracts are traded. Individual options are indexed by $n = 1, \ldots, N$ according to the following characteristics: strike price, expiration date, and put/call identifier. We recognize that the main determinants of option price changes are the changes in the value of the underlying and its volatility (e.g., Bakshi et al. (1997)).

The option price change can therefore be written as:

$$d^n_{t+1} = \Delta_{n,t+1} d^n_t + v_{n,t+1} v^n_t + \epsilon_{n,t+1},$$

(1)

where $d^n_{t+1} = P_{n,t+1} - P_{n,t}$ is the price change of option $n$ at time $t+1$, $\Delta_{n,t}$ and $v_{n,t}$ are the delta and vega of the option, respectively, and $d^n$ and $v^n$ denote changes in the price of the underlying and its volatility, i.e., $d^n_t = P^n_t - P^n_{t-1}$ and $d^n_t = \nu_t - \nu_{t-1}$. The error $\epsilon_{n,t+1}$ contains terms of order $dt$ of the option pricing model. We define vega as the first partial derivative of the option price with respect to volatility. One of the key challenges is that volatility is not directly observable, and therefore needs to be estimated from market data. We provide further details on the link between volatility and observable volatility indices in Section 3.2.

We disaggregate an option trade into its exposure to the underlying asset (through the option delta) and to its volatility (through the option vega). The key advantage of this linear transformation is that option order flows are now expressed in the same units, and can be meaningfully aggregated across options. We define the time-$t$ aggregate net dollar exposure to the underlying by $x_t^d$ and to the volatility by $x_t^v$:

$$x_t^d = \sum_i q_i \times \Delta_{i,t} \times BuySell_i \times P^n_{i,t-1},$$

(2)

$$x_t^v = \sum_i q_i \times v_{i,t} \times BuySell_i \times \nu_{t-1}.$$
We sum over all option trades (indexed by \(i\)) in the fixed time interval between \(t - 1\) and \(t\). With a slight abuse of notation, we denote by \(A_{ij}\) and \(v_{ij}\) the delta and vega of the particular option trade \(i\) prevailing at the start of interval (time \(t - 1\)). Further, \(Q_j\) is the trade volume in number of options, and \(BuySell_j\), a binary variable that equals 1 for a buyer originated trade and \(-1\) for a seller originated trade. In the remainder of the paper, \(x^2_t\) and \(x^4_t\) are called delta order flow and vega order flow, respectively. The delta order flow is denoted in U.S. dollars, which allows for a meaningful comparison to the order flow in the underlying asset.

In the next step, we relate the delta and vega order flows to the price changes of the underlying (\(d^n_t\)) and volatility (\(d^v_t\)) in a VAR framework. We include the dollar order flow of the SPY ETF (denoted \(x^{pp}_t\)) as an additional equation, as it is a proxy for the trading volume in the underlying asset. This approach also accounts for trading strategies that involve simultaneous trading in options and the underlying. Note that the order flow \(x^{pp}_t\) is measured in the same unit as \(x^2_t\), as both represent dollar order flow exposure to the underlying. The difference, however, is that the latter is constructed from option order flows only.

The empirical model can be written as follows:

\[
\begin{align*}
\Delta d^n_t &= c_1 + \sum_{i=1}^{6} A_{1i} \Delta d^n_{i,t-1} + \sum_{i=1}^{6} B_{1i} \Delta d^v_{i,t-1} + \sum_{i=1}^{6} C_{1i} \Delta x^{pp}_{i,t-1} + \sum_{i=1}^{6} D_{1i} x^2_{i,t-1} + \sum_{i=1}^{6} E_{1i} x^4_{i,t-1} + \epsilon_{1,t} \\
\Delta d^v_t &= c_2 + \sum_{i=1}^{6} A_{2i} \Delta d^n_{i,t-1} + \sum_{i=1}^{6} B_{2i} \Delta d^v_{i,t-1} + \sum_{i=1}^{6} C_{2i} \Delta x^{pp}_{i,t-1} + \sum_{i=1}^{6} D_{2i} x^2_{i,t-1} + \sum_{i=1}^{6} E_{2i} x^4_{i,t-1} + \epsilon_{2,t} \\
\Delta x^{pp}_t &= c_3 + \sum_{i=1}^{6} A_{3i} \Delta d^n_{i,t-1} + \sum_{i=1}^{6} B_{3i} \Delta d^v_{i,t-1} + \sum_{i=1}^{6} C_{3i} \Delta x^{pp}_{i,t-1} + \sum_{i=1}^{6} D_{3i} x^2_{i,t-1} + \sum_{i=1}^{6} E_{3i} x^4_{i,t-1} + \epsilon_{3,t} \\
\Delta x^2_t &= c_4 + \sum_{i=1}^{6} A_{4i} \Delta d^n_{i,t-1} + \sum_{i=1}^{6} B_{4i} \Delta d^v_{i,t-1} + \sum_{i=1}^{6} C_{4i} \Delta x^{pp}_{i,t-1} + \sum_{i=1}^{6} D_{4i} x^2_{i,t-1} + \sum_{i=1}^{6} E_{4i} x^4_{i,t-1} + \epsilon_{4,t} \\
\Delta x^4_t &= c_5 + \sum_{i=1}^{6} A_{5i} \Delta d^n_{i,t-1} + \sum_{i=1}^{6} B_{5i} \Delta d^v_{i,t-1} + \sum_{i=1}^{6} C_{5i} \Delta x^{pp}_{i,t-1} + \sum_{i=1}^{6} D_{5i} x^2_{i,t-1} + \sum_{i=1}^{6} E_{5i} x^4_{i,t-1} + \epsilon_{5,t},
\end{align*}
\]

where \(A_{1i}, \ldots, E_{ij}\) are constant coefficients. Eq. (4) emphasizes the ordering of the equations in the structural VAR. This ordering identifies the orthogonal structural innovations, as it restricts the contemporaneous innovation in one variable to affect that of another, but not the reverse.\(^6\) In particular, in the first equation of the system, price changes \(d^n_t\) are affected by all other variables contemporaneously but do not affect any other variables. In the last equation, the vega order flow \(x^4_t\) contemporaneously affects all other variables but is not affected by others.

While in general the ordering is arbitrary, we believe that the model in Eq. (4) is the most natural choice. First, we follow Hasbrouck (1991) by placing price changes first and order flows second. This is motivated by sequential trade models, where order flows cause returns because of asymmetric information and informed trading, whereas returns do not directly cause order flows. With regard to the ordering of the price and volatility changes \(d^n_t\) and \(d^v_t\), we note that the former is a traded asset whereas the latter (the volatility) is not. Being a traded asset, \(d^n_t\) should respond fast and efficiently to new information, such that it is affected by contemporaneous innovations in volatility. The reverse does not hold, since volatility moves slower as it is only indirectly traded through a wide range of options. We also believe our model represents the most conservative ordering. Our focus is on measuring the information content in option order flows after controlling for any information captured by trading in the underlying itself. Accordingly, our proposed ordering attributes any contemporaneous information captured in, for example, both \(x^{pp}_t\) and \(x^2_t\) to the former when predicting \(d^n_t\). While this ordering may create a bias by reducing the explanatory power of \(x^2_t\) on \(d^n_t\), it is the most conservative approach for our purposes. A number of alternative specifications are studied in the robustness section.

The proposed model yields several practical advantages. First, the lags in the VAR system naturally account for any autocorrelation in the volatility process. Second, with separate equations for the underlying and volatility, we obtain price impact estimates in both components. Third, the analysis of option markets typically requires a data reduction technique as the number of actively traded options is large (considering put and call options, as well as different strike prices and expiration dates).\(^7\) To reduce the dimensionality of the data, some papers suggest modeling option prices as a function of time to maturity and moneyness (e.g., Alfa-Sahalia and Lo (2000)). In this paper, we apply a data reduction motivated by the two-factor option pricing model of Heston (1993).

3. Empirical setup

3.1. Data

Our trade and quote data are from the Refinitiv Tick History database.\(^8\) The data set consists of European-style SPX options written on the S&P 500 Index and covers the time period from January 2, 2014 to December 31, 2014. Option contracts differ in strike price, put/call identifier, and expiration date and are part of the third-Friday expiration cycle, which includes the most liquid

\(^{6}\) A change in the value of the underlying changes the option delta (and vega), which may create a spurious correlation between delta flow (and vega flow) and the underlying. By using the delta and vega at the start of the interval, we prevent this mechanical relation.

\(^{7}\) This approach is equivalent to estimating the reduced form VAR and applying a Cholesky decomposition on the residuals [see Chapter 9.4 in Hasbrouck (2007)].

\(^{8}\) Chan et al. (2002) also apply an extension of the Hasbrouck (1991) model to incorporate option markets in their study, but use only data of one put and call option in their empirical analysis.

\(^{9}\) The database used to operate under the name Thomson Reuters Tick History database.
option contracts written on the S&P 500 Index. For each option trade, we observe the exact time of the trade (to the millisecond), as well as the trade price and volume. Quote data consist of the best bid and ask prices, which are used to classify each trade as a market buy or sell order.\textsuperscript{10} We also collect corresponding trade and quote data for the SPY ETF, which we use as a tradeable proxy for the underlying asset and to construct the price changes $d_t$.

We apply a range of standard filters to our data. First, we discard options with less than seven or more than 180 calendar days to maturity. Long-term options are very infrequently traded, and short-term options are adversely affected by changes in the expiration cycle. Second, we remove individual trades for which the dollar volume exceeds $10$ million, as these are likely negotiated over the counter (OTC).\textsuperscript{11} And third, we avoid opening and closing auctions by restricting our analysis to trades between 8:45 am to 2:45 pm CT, leaving a sample of six full trading hours.

3.2. Model design choices and implementation

The empirical implementation of the proposed model requires several design choices, such as the sample frequency and the construction of the order flow variables $x_t^2$ and $x_t^v$, as well as volatility changes $d_t^v$. We construct these as follows.

**Sampling frequency:** We sample data at the hourly frequency. While our data set and methodological framework allows us to conduct the empirical analysis at even higher frequencies, a careful analysis in Section 5 shows that the hourly frequency provides the best compromise between a high-frequency analysis and avoiding problems related to microstructure noise in option prices. That is, option prices are characterized by relatively large bid–ask spreads and tick sizes, which imply that prices adjust infrequently—only after sufficiently large changes in the underlying or volatility. Further, the volatility process is not directly observed and must be estimated, and the resulting measurement error becomes problematic at too high sample frequencies.

**Delta and vega:** A second important design choice relates to the calculation of trade deltas and vegas, which are required for the order flow variables in Eqs. (2) and (3). We use the smile-consistent option pricing model of Heston (1993) to calculate deltas and vegas for each individual trade. Our methodology can be summarized as follows. First, we use the theoretical functional relationship between the (observable) VIX index and the (unobservable) spot volatility to simultaneously calibrate the Heston model to intra-daily option data and the VIX index on the first day of our sample. Our approach avoids filtering techniques and high-dimensional optimization (Broadie et al., 2007; Christoffersen et al., 2010) and is easy to extend to other pricing models or option markets. Second, we use the calibrated parameters from the first day of the sample and VIX index values on the next trading day to obtain the out-of-sample S&P 500 spot volatility. Third, equipped with estimates of the spot volatility and the structural parameters of the model, we then calculate deltas and vegas for all option trades on the second day using standard Fourier inversion techniques. The procedure is then repeated for all trading days in our sample. Our methodology allows for a conceptually easy construction of the trading exposure to the underlying S&P 500 ($x^2$) and the S&P 500 volatility ($x^v$) using the definitions in Eqs. (2) and (3). We provide a more detailed description of our methodology in Appendix.

**Volatility process:** We use the VIX index values to estimate the volatility process of the S&P 500 at the one minute frequency.\textsuperscript{12} We follow a standard procedure, and define the time-$\tau$ VIX index with maturity $\tau$ by:

$$VIX(\tau) = \left[2e^{\tau} \int_0^\infty \frac{O(t, k, t+\tau)}{k^2} dk \right]^{1/2},$$

where $O(t, K, T)$ denotes the time-$t$ quoted midpoint of an out-of-the-money (OTM) option (with strike $K$ and maturity $T$).\textsuperscript{13} To approximate the integral in Eq. (5), we first construct an option pricing function that is continuous in the strike $K$. We follow previous studies and interpolate the implied volatilities of OTM options by a cubic polynomial and extrapolate the curve by fixing the implied volatilities of options beyond the traded strike range to the nearest available market implied volatility [for these procedures, see Broadie et al. (2007) and Carr and Wu (2009)]. We then use a simple adaptive Gauss–Kronrod quadrature method to calculate the integral in Eq. (5). In the Appendix, we describe the link between the VIX and spot volatility in more detail.

4. Results

4.1. Summary statistics

**Option bid–ask spreads:**

We first show implicit trading costs in terms of option bid–ask spreads and next compare these to estimated price impacts. Table 2 contains detailed spread results for our sample (including those reported in Table 1). The nature of option markets is such that percentage and dollar spreads are not easily comparable between options with different degrees of moneyness. In particular, deep out-of-the-money options have a low value and therefore the percentage spread is typically large compared to its dollar spread. The opposite holds for deep in-the-money options. Further, we show both value-weighted and equal-weighted spreads, because the former overweight high-priced ITM options and the latter overweight low-priced OTM options.

---

\textsuperscript{10} There have been some issues with signing option trades using Lee–Ready because trades can occur within the quoted prices and are often timed strategically (see Muravyev and Pearson (2020)). This adds measurement error to the order flow variables, but will otherwise not affect the analyses.

\textsuperscript{11} This filter drops less than 0.09% of the trades.

\textsuperscript{12} The higher one-minute frequency allows for more precise delta and vega estimates used for each option trade.

\textsuperscript{13} Moneyness is defined relative to the forward price of the underlying.
Table 2
Bid-ask spreads of SPX options.

<table>
<thead>
<tr>
<th>Variable</th>
<th>All options</th>
<th>In-the-money</th>
<th>At-the-money</th>
<th>Out-of-the-money</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean</td>
<td>sdev</td>
<td>mean</td>
<td>sdev</td>
</tr>
<tr>
<td><strong>Panel A. Equally weighted</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effective quoted half spread (%)</td>
<td>12.81</td>
<td>17.14</td>
<td>3.25</td>
<td>3.37</td>
</tr>
<tr>
<td>Effective spread (%)</td>
<td>5.60</td>
<td>13.07</td>
<td>0.93</td>
<td>1.77</td>
</tr>
<tr>
<td>Effective spread MP (%)</td>
<td>5.58</td>
<td>15.64</td>
<td>1.25</td>
<td>36.54</td>
</tr>
<tr>
<td>Realized spread 10 m (%)</td>
<td>2.53</td>
<td>14.46</td>
<td>0.25</td>
<td>5.28</td>
</tr>
<tr>
<td>Adverse selection 10 m (%)</td>
<td>3.12</td>
<td>15.46</td>
<td>0.69</td>
<td>5.28</td>
</tr>
<tr>
<td>Realized spread 30 m (%)</td>
<td>2.24</td>
<td>17.31</td>
<td>0.22</td>
<td>8.61</td>
</tr>
<tr>
<td>Adverse selection 30 m (%)</td>
<td>3.38</td>
<td>18.45</td>
<td>0.71</td>
<td>8.60</td>
</tr>
<tr>
<td>Effective quoted half spread ($)</td>
<td>0.60</td>
<td>0.62</td>
<td>1.35</td>
<td>0.72</td>
</tr>
<tr>
<td>Effective spread ($)</td>
<td>0.20</td>
<td>0.37</td>
<td>0.34</td>
<td>0.45</td>
</tr>
<tr>
<td>Realized spread 10 m ($)</td>
<td>0.13</td>
<td>0.78</td>
<td>0.10</td>
<td>1.73</td>
</tr>
<tr>
<td>Adverse selection 10 m ($)</td>
<td>0.08</td>
<td>0.70</td>
<td>0.24</td>
<td>1.73</td>
</tr>
<tr>
<td>Realized spread 30 m ($)</td>
<td>0.12</td>
<td>1.21</td>
<td>0.09</td>
<td>2.81</td>
</tr>
<tr>
<td>Adverse selection 30 m ($)</td>
<td>0.08</td>
<td>1.16</td>
<td>0.26</td>
<td>2.81</td>
</tr>
<tr>
<td><strong>Panel B. Dollar volume weighted</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effective quoted half spread (%)</td>
<td>3.27</td>
<td>3.55</td>
<td>1.29</td>
<td>1.90</td>
</tr>
<tr>
<td>Effective spread (%)</td>
<td>1.69</td>
<td>2.78</td>
<td>0.61</td>
<td>1.35</td>
</tr>
<tr>
<td>Realized spread 10 m (%)</td>
<td>1.63</td>
<td>4.19</td>
<td>0.58</td>
<td>2.64</td>
</tr>
<tr>
<td>Adverse selection 10 m (%)</td>
<td>0.05</td>
<td>3.37</td>
<td>0.04</td>
<td>2.35</td>
</tr>
<tr>
<td>Realized spread 30 m (%)</td>
<td>1.66</td>
<td>6.02</td>
<td>0.59</td>
<td>4.09</td>
</tr>
<tr>
<td>Adverse selection 30 m (%)</td>
<td>0.03</td>
<td>5.48</td>
<td>0.03</td>
<td>3.90</td>
</tr>
<tr>
<td>Quoted half spread ($)</td>
<td>1.21</td>
<td>0.85</td>
<td>1.49</td>
<td>0.69</td>
</tr>
<tr>
<td>Effective spread ($)</td>
<td>0.59</td>
<td>0.67</td>
<td>0.72</td>
<td>0.76</td>
</tr>
<tr>
<td>Realized spread 10 m ($)</td>
<td>0.59</td>
<td>1.22</td>
<td>0.70</td>
<td>1.77</td>
</tr>
<tr>
<td>Adverse selection 10 m ($)</td>
<td>0.01</td>
<td>1.01</td>
<td>0.03</td>
<td>1.57</td>
</tr>
<tr>
<td>Realized spread 30 m ($)</td>
<td>0.60</td>
<td>1.87</td>
<td>0.62</td>
<td>2.83</td>
</tr>
<tr>
<td>Adverse selection 30 m ($)</td>
<td>0.00</td>
<td>1.73</td>
<td>0.11</td>
<td>2.69</td>
</tr>
</tbody>
</table>

This table shows various quote and trade-based spreads of all SPX options in 2014. We show statistics for the full sample, as well as for subsamples of in-the-money ($|d| > 0.65), at-the-money (0.65 > |d| > 0.35), and out-of-the-money (0.35 > |d|), with |d| denoting the absolute value of the option delta. Panel A shows equally-weighted observations and Panel B the dollar volume-weighted observations (price times quantity). We define the effective quoted half-spread as the bid–ask spread prevailing just before a trade. The effective half-spread is defined as \( (p - m_{\mid \tau \mid}) \), which is then decomposed into a realized spread \( (p - m_{\mid \tau \mid}) \), and adverse selection component \( (m_{\mid \tau \mid} - m_{\mid \tau \mid}) \), based on a midpoint price \( \tau \) minutes later. We also report the Muravyev–Pearson corrected effective spread (MP), which replaces the midpoint by the predicted value of a regression of the midpoint on Black–Scholes price minus midpoint, delta times lagged underlying price differences, and lagged price changes. All variables are reported in dollars and in percentages as a fraction of the midpoint.

The equal-weighted average effective quoted half spread, i.e., the quoted half-spread just before each trade, is 12.8%, while the effective spread is only 5.6%. The large difference is explained by the many trades negotiated off-exchange or directly with dealers that occur inside the quoted bid and ask prices. The dollar value-weighted percentage spreads are about one-third of the equal-weighted values, reflecting that high-priced ITM options have relatively small percentage spreads. We further see that the sample average spread reduces to 5.58% after applying the correction of Muravyev and Pearson (2020), who show that the use of the midpoint in the effective spread calculation is inappropriate when investors are more likely to buy (sell) when the unobserved fundamental is closer to the ask (bid). The correction does not change results as much for SPX options as it does for equity options, likely because many SPX trades already occur within the quoted bid and ask prices.

The table also shows the effective spread in the ten- or thirty-minute realized spread and adverse selection components. The equal-weighted adverse selection component is large, 3.12% at the ten-minute level, compared to a realized spread of 2.53%. In contrast, the dollar volume-weighted average is only 0.05%, compared to the realized spread of 1.66%. This difference is mainly caused by deep OTM option trades, which have low dollar volumes and relatively large adverse selection components.

**Individual option trades:**

Table 3 reports summary statistics of individual option trades. The full sample consists of more than 1.8 million trades. Column (1) shows that trades are generally large with an average size of $93,300, which is about 20 times larger than a typical stock or ETF trade. Note that the dollar risk that changes hands in these trades is even larger, considering the embedded leverage options offered. The average SPX option trade is 51.4 contracts, and each contract has a multiplier equal to 100. Column (1) further shows that 41% of the option trades are calls (hence 59% are puts) and that 66.2% of the trades are buyer originated (therefore 33.8% are seller originated trades). The relatively strong buying demand in put options suggests that investors use options to hedge the downside risk in the S&P 500 Index.

---

14 They replace the midpoint by the predicted value of a regression of the midpoint on Black–Scholes price minus midpoint, delta times lagged underlying price differences, and lagged price changes. This adjustment offers a better expectation of the unobserved fundamental value than the quoted midpoint.
The table shows the mean and standard deviation (in parentheses) of option trade variables. The first column reports results for the full sample, and the remaining columns those for nine subsets of the data double sorted by moneyness and time to maturity. Each subsets can be identified by the rows showing the average Moneyness and Time to Maturity. Most variables are self explanatory. The vegas are calculated with respect to the underlying volatility. The trade direction equals 1 if the trade is originated with a market buy order, and −1 if it is a market sell order. The call indicator equals 1 if the trade is in a call option, and zero if it is in a put option.

The absolute delta of put and call options is on average 0.248 with a standard deviation of 0.244, which suggests that there is considerable variation in the exposure option trades exhibit with respect to the underlying. Option trades also vary substantially in their volatility exposure with an average vega of 73.1 and a standard deviation of 50.2. This implies that an increase in (annualized) volatility from 15% to 16% leads to an increase in the price of a typical option by approximately $0.73. Overall, these numbers indicate that investors actively trade options with a wide variety of characteristics, presumably to obtain varying exposures to the underlying and volatility. Column (1) also reveals that out-of-the-money options are traded relatively frequently, as the average moneyness (0.945) is less than one. Moneyness is defined as $F_{t,T}/K$ for puts and $K/F_{t,T}$ for calls where $F_{t,T}$ denotes the time-$t$ forward price of the underlying. The average time to maturity is 0.118 years, or about 30 trading days.

The remaining columns in Table 3 report summary statistics for nine subsets of the data, double sorted by time to maturity (ttm) and moneyness. We create three buckets by ttm using the cutoff values of 30 and 90 days. We use cutoff values of 0.95 and 1 to group options into either deep out-of-the-money, out-of-the-money or in-the-money. While these cutoff values are arbitrary, this choice ensures that each subset includes a reasonable number of trades. The same subsamples are used throughout the remainder of the analyses. Each subset can be identified in the table by the corresponding average moneyness and ttm. As expected, the subgroups that contain near-the-money options yield the highest exposure to volatility (those in columns (3), (6), and (9)). Similarly, the subsets with highest moneyness, in columns (4), (7), and (10), yield the highest exposure to the underlying.

**Aggregated order flow variables:**

Table 4 provides the summary statistics of all variables used in the estimation of Eq. (4), including hourly order flow variables (in exposures of $100 million), as well as the changes in the price of the underlying asset and its volatility. Our sample period is characterized by selling pressure in the SPY ETF, as the average of the order flow $x_{SPY}$ is $-43$ million per hour (with a standard deviation of $432$ million). The delta order flow constructed from option trades, measured in the same units, is positive with $179$...

---

15 These subsets contain deep in-the-money calls (with a delta close to one) and deep in-the-money puts (with a delta just below zero). This explains why the average absolute delta of trades in these subsets ranges between 0.5 and 0.6.

16 This unit is most convenient when interpreting the structural VAR coefficients.
Table 5
Long-run impulse response functions VAR model.

<table>
<thead>
<tr>
<th>Response</th>
<th>Impulse</th>
<th>σ(1)</th>
<th>d^0</th>
<th>d^1</th>
<th>x^0</th>
<th>x^1</th>
<th>x^2</th>
<th>x^3</th>
<th>x^4</th>
<th>x^5</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>d^0 (cents)</td>
<td>189.93</td>
<td>215.70***</td>
<td>−277.10***</td>
<td>171.00***</td>
<td>98.13***</td>
<td>−55.64***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(11.28)</td>
<td>(14.14)</td>
<td>(8.06)</td>
<td>(4.41)</td>
<td>(2.59)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d^0 (percentage points)</td>
<td>0.50</td>
<td>0.00</td>
<td>0.52***</td>
<td>−0.23***</td>
<td>−0.11***</td>
<td>0.07***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.09)</td>
<td>(18.75)</td>
<td>(7.45)</td>
<td>(3.31)</td>
<td>(2.36)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>x^0 ($100 million)</td>
<td>4.20</td>
<td>−0.38*</td>
<td>0.62***</td>
<td>4.91***</td>
<td>0.44</td>
<td>−0.28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.67)</td>
<td>(2.86)</td>
<td>(19.99)</td>
<td>(1.62)</td>
<td>(1.10)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>x^1 ($100 million)</td>
<td>8.23</td>
<td>2.46***</td>
<td>−1.28***</td>
<td>1.03**</td>
<td>9.87***</td>
<td>−0.16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(5.17)</td>
<td>(2.75)</td>
<td>(2.11)</td>
<td>(18.21)</td>
<td>(0.29)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>x^2 (percentage points)</td>
<td>0.16</td>
<td>0.01</td>
<td>0.02**</td>
<td>−0.02*</td>
<td>−0.01</td>
<td>0.21***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.62)</td>
<td>(1.99)</td>
<td>(1.73)</td>
<td>(0.51)</td>
<td>(20.60)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This table shows the long-run effects of the impulse response functions (IRFs) of the VAR model estimated by Eq. (4). Long-run is defined as four periods (trading hours). The first column shows the standard deviation of the structural residual of each equation, which represents the size of the shocks in the IRF. The next columns show the long-run impact of a one standard deviation impulse to each of the variables. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 5 offers two main results. First, column (6) shows that a one standard deviation shock in \( x^1 \) increases \( d^0 \) by 0.07. This translates into a volatility increase from the sample average of 11.5% to 11.57%, and corresponds to 0.14 standard deviations of \$864 million per hour, and is twice as volatile with a standard deviation of \$16 million.\(^\text{17}\) This finding is in line with the summary statistics for individual trades, and confirms that investors in the SPX market trade in size.

The risk that changes hands through a one standard deviation delta order flow trade is about twice as large as that of a vega order flow trade. To see this, note that the standard deviation of delta order flow is much larger than the standard deviation of vega order flow (\$864M vs. \$16M). However, as a pricing component, the volatility returns are much riskier than the underlying returns.\(^\text{18}\) Together, the hourly risk transfer of a one standard deviation delta flow shock for one hour is \( 864 \times 0.2\% = 1.75 \) million, while that of a vega order flow trade is \( 16 \times 5\% = 0.8 \) million.

### 4.2. Price impact model

We estimate the VAR model in Eq. (4) with two lags as suggested by the Akaike information criterion. Following Hasbrouck (1991), we assume the trading process restarts at the beginning of each trading day, and therefore we set all initial lagged values to zero.

Table 5 provides the long run impulse response functions (IRFs) of the model. The table contains the five-by-five matrix of IRFs, which shows the impact of an impulse in one variable (in columns) on the cumulative impact on all other variables (in rows). We allow each shock to iterate through the system for four periods. The table also shows the size of each shock in the first column.

### 4.2.1. Main results

Table 5 offers two main results. First, column (6) shows that a one standard deviation shock in \( x^1 \) increases \( d^0 \) by 0.07. This translates into a volatility increase from the sample average of 11.5% to 11.57%, and corresponds to 0.14 standard deviations of \$864 million per hour, and is twice as volatile with a standard deviation of \$16 million.\(^\text{17}\) This finding is in line with the summary statistics for individual trades, and confirms that investors in the SPX market trade in size.

The risk that changes hands through a one standard deviation delta order flow trade is about twice as large as that of a vega order flow trade. To see this, note that the standard deviation of delta order flow is much larger than the standard deviation of vega order flow (\$864M vs. \$16M). However, as a pricing component, the volatility returns are much riskier than the underlying returns.\(^\text{18}\) Together, the hourly risk transfer of a one standard deviation delta flow shock for one hour is \( 864 \times 0.2\% = 1.75 \) million, while that of a vega order flow trade is \( 16 \times 5\% = 0.8 \) million.

\(^{17}\) The hourly standard deviation of \$864 delta flow can be constructed from the individual option trade data in Table 3 as follows. The average option trade has a delta flow of about \$2.5 million (a delta of 0.25 times 51 contracts times a 100 multiplier, times 1930 (value underlying)). A given trading hour has on average about a 1000 trades, so if the hourly standard deviation of net flow (buys minus sells) is about 40%, then the standard deviation of hourly delta flow is \$1,000 million (40% of 1000 times 2.5 million), which is fairly close to the reported standard deviation of \$823 million.

\(^{18}\) The annualized volatility level is on average 11.5% and its difference has a standard deviation of 0.58 percentage points, which gives a volatility return standard deviation of 5.0% = 0.58/11.5 × 100. The underlying has a spot price of \$1,930 on average and its difference has a standard deviation of \$3.92 per hour, giving a return standard deviation of 0.2% = 3.92/1930 × 100.
This magnitude can be interpreted as the price impact of volatility speculation, and is economically relevant despite very low levels of volatility in 2014. However, this impact results from a one standard deviation vega-flow shock based on the aggregate SPX option market, which is a massive transfer of risk with an hourly return standard deviation of $0.8$ million (see Footnote 19). Related studies typically analyze the volatility impact of trading in one or two options (e.g., Bollen and Whaley (2004), Rourke (2013)), but these results are not representative of informed trading in the general market and may be biased by not adjusting for cross-option trading strategies. The remaining variation in volatility is driven by public information arrival and the other order flows.

The second main result is shown in column (5), where the long-run impact of a shock in $x^d$ is $98.13$ cents. This can be interpreted as a change in the average value of the underlying from $1,930$ to $1,930.98$, which is an increase of $5.1$ bps. This price impact is very small compared to the average option effective spread of $169$ bps. Further, even though $0.98$ represents about $25\%$ of the hourly standard deviation ($3.92$) of changes in the underlying, note that it requires a massive order flow exposure shock of $823$ million (Table 5). This makes the dollar price impact very small. We do note that these low price impacts in SPX options are consistent with the low but positive price impacts for equity options in Muravyev (2016). Interestingly, the impact of an order flow shock in the ETF ($x^{SPY}$) is larger with $1.71$ per standard deviation (see column (4)). Also note that the size of ETF shocks is much smaller than the size of delta flow shocks: $420$ million vs. $823$ million. Within the context of our model, this would suggest that the per-dollar price impact of an ETF shock is $3.4$ times larger than when the same exposure is obtained through options. Of course, our model does not include order flows in other (near-perfectly) correlated assets like the E-mini futures, and thus does not control for any correlation with those order flows.

Table 5 provides several additional findings. The impact of $x^{SPY}$ on $d^u$ is negative and large, and in Section 4.3 we show this is fully explained by the leverage effect (i.e., the negative correlation between $d^u$ and $d^u$). The mechanism is that $x^{SPY}$ increases $d^u$ and simultaneously decreases $d^u$; this contemporaneous correlation cannot be disentangled by the VAR model. The same argument explains why $x^d$ negatively affects $d^u$, and $x^d$ negatively affects $d^d$. Further, the coefficients on the diagonal of Table 5 reveal the long-term impact of a shock on the variable itself. All variables are positively autocorrelated, because the long-term impacts are greater than the size of the structural shocks (shown in the first column). In addition, $x^d$ and $x^u$ are uncorrelated in the long-run as a shock in one does not affect the other. Indeed, $x^d$ and $x^u$ mechanically have a positive correlation for call option trades and a negative correlation for put option trades — on average, the two opposing effects seem to cancel out. Lastly, ETF flow $x^{SPY}$ significantly causes delta flow in the long-run, but not the other way around. This suggests that $x^{SPY}$ is leading and delta flow is following.

Fig. 1 shows the cumulative impulse response functions over time. For brevity, we only show the shocks of the order flow variables on the changes in the price and volatility. The figure reveals that option markets are very efficient, in the sense that the majority of order flow information is reflected in prices contemporaneously.

Table 6 contains the regression coefficients of the VAR model estimated by OLS. The omitted coefficients in each column reveal the ordering of the equations in the structural model, which in turn identifies the structural residuals. The results of the IRFs are clearly visible in the coefficients as well. Of note is the R-squared of $d^u$ of $0.25$, meaning a quarter of its variation can be explained by the three contemporaneous order flows and all the lagged variables. The R-squared of $d^u$ in column (5) is very high with $0.766$, but this is largely driven by $d^u$, which appears contemporaneously in the regression and picks up the leverage effect (with a t-stat of 32.2). Furthermore, the positive contemporaneous coefficient of $x^d$ on $x^{SPY}$ (0.086) suggests that either traders obtain exposure to the underlying by trading simultaneously options and the ETF, or that market makers in the option market hedge their position in the underlying (as in Hu (2018)).

4.2.2. Economic implications

Individual and option portfolio price impact. An option trade affects the underlying and volatility through its delta and vega flow, which in turn affects the option price because it linearly depends on the underlying and volatility. We now calculate these option price impacts using the estimated VAR results for several hypothetical trades in individual options and option portfolios. These price impacts are a component of total trading costs and can be meaningfully compared to the bid–ask spreads of Table 2. Table 7 shows the numerical results for hypothetical portfolio trades based on data of June 27, 2014 at 12:00 (the middle of our sample period). The first rows correspond to a straddle, with a long ATM call and put. We consider a very large trade of 1000 contracts [about 20 times the sample average—see column (6) in Table 3], where each contract has a multiplier of 100, at a price of $16.88$ (call) and $16.90$ (put) per option. The dollar cost is $1.69$ million for each leg, but due to the embedded leverage, the delta flow is $124$ million for the call and $74$ million for the put. These massive exposures generate only a tiny price impact

---

Footnotes:

19 For example, if investors frequently trade straddles (i.e., a long put and call with similar strike price and expiration date) then the vega exposures of the order flow are positively correlated across option contracts. This leads to an omitted variables bias that inflates coefficients when the analysis does not include the order flows of all contracts.

20 Given that these estimates stem from an SVAR model, the technically appropriate number to look at is the forecast-error variance decomposition: the total variation in the structural SPY innovations explained by the different order flows, which account for the ordering of the equations. These numbers are 7.6% for delta-flow $x^d$, 22.5% for SPY flow $x^{SPY}$; 1.4% for vega flow $x^u$; and 45% for the volatility process $d^u$ (this large value represents the leverage effect), leaving 23.5% of non-trade related (external) variation.

21 The intuition is that $x^d$ positively predicts $d^u$, and therefore negatively predicts $d^u$ because of the negative correlation between $d^u$ and $d^u$. Similarly, $x^d$ positively predicts $d^u$, and therefore negatively predicts $d^u$.

22 Indeed, the underlying and volatility process are simultaneously determined and have a strong negative correlation of $-0.858$ at the hourly frequency.

23 A single ATM call option, at a price of only $16.88$, offers a linearized exposure of about $1,200$ to the underlying, because its delta is $0.63$ relative to the value of the SPY ETF of $1,954$. 
At option portfolio trades, this price impact may shrink significantly depending on the extent that exposures to the underlying and volatility.

Second, when looking at put and call spreads the price impact of the individual option trades – thus ignoring cross-option price impact of the straddle – and these are 29.9 bps for the call, 42.1 bps for the put, yielding a value weighted portfolio average of 36 bps. As the price impact is linear, the impact for the underlying of $0.104 and $-0.131 dollars for the call and put, respectively.

We next convert these impacts on the underlying and volatility to obtain the price impact in the option. Column (10) shows the price impact of the individual option trades – thus ignoring cross-option price impact of the straddle – and these are 29.9 bps for the call, 42.1 bps for the put, yielding a value weighted portfolio average of 36 bps.

Table 6
Regression coefficients VAR model.

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>x*</td>
<td>-</td>
<td>0.039</td>
<td>-2.137**</td>
<td>0.457***</td>
<td>37.547</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0)</td>
<td>(-2.1)</td>
<td>(3.9)</td>
<td>(1.0)</td>
</tr>
<tr>
<td>L1 x*</td>
<td>0.205***</td>
<td>0.737</td>
<td>0.410</td>
<td>-0.120</td>
<td>-57.625</td>
</tr>
<tr>
<td></td>
<td>(3.1)</td>
<td>(0.4)</td>
<td>(0.5)</td>
<td>(-0.9)</td>
<td>(-1.5)</td>
</tr>
<tr>
<td>L2 x*</td>
<td>0.0266</td>
<td>0.0355</td>
<td>0.131</td>
<td>-43.673</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.7)</td>
<td>(0.0)</td>
<td>(0.2)</td>
<td>(-0.8)</td>
<td>(-1.1)</td>
</tr>
<tr>
<td>x*</td>
<td>-</td>
<td>-</td>
<td>0.086**</td>
<td>-0.012**</td>
<td>2.735***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(6.0)</td>
<td>(-6.8)</td>
<td>(3.8)</td>
</tr>
<tr>
<td>L1 x*</td>
<td>0.001</td>
<td>0.178***</td>
<td>-0.031**</td>
<td>0.002</td>
<td>0.446</td>
</tr>
<tr>
<td></td>
<td>(0.5)</td>
<td>(4.3)</td>
<td>(-2.0)</td>
<td>(0.7)</td>
<td>(0.6)</td>
</tr>
<tr>
<td>L2 x*</td>
<td>-0.001</td>
<td>-0.076***</td>
<td>0.001</td>
<td>0.003</td>
<td>0.263</td>
</tr>
<tr>
<td></td>
<td>(-0.8)</td>
<td>(-2.3)</td>
<td>(0.0)</td>
<td>(1.5)</td>
<td>(0.4)</td>
</tr>
<tr>
<td>x**</td>
<td>-</td>
<td>-</td>
<td>-0.056***</td>
<td>14.799***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(-10.6)</td>
<td>(8.5)</td>
<td></td>
</tr>
<tr>
<td>L1 x**</td>
<td>-0.002</td>
<td>-0.002</td>
<td>0.165***</td>
<td>0.011***</td>
<td>-1.441</td>
</tr>
<tr>
<td></td>
<td>(-1.3)</td>
<td>(-0.0)</td>
<td>(4.0)</td>
<td>(2.7)</td>
<td>(-1.0)</td>
</tr>
<tr>
<td>L2 x**</td>
<td>0.000</td>
<td>-0.021</td>
<td>0.040</td>
<td>0.003</td>
<td>-5.321***</td>
</tr>
<tr>
<td></td>
<td>(0.1)</td>
<td>(-0.4)</td>
<td>(1.1)</td>
<td>(0.9)</td>
<td>(-3.6)</td>
</tr>
<tr>
<td>d*</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-525.690***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(-32.2)</td>
<td></td>
</tr>
<tr>
<td>L1 d*</td>
<td>0.028</td>
<td>1.665**</td>
<td>-0.176</td>
<td>0.021</td>
<td>16.631</td>
</tr>
<tr>
<td></td>
<td>(1.3)</td>
<td>(2.1)</td>
<td>(-0.4)</td>
<td>(0.4)</td>
<td>(0.8)</td>
</tr>
<tr>
<td>L2 d*</td>
<td>0.012</td>
<td>1.566</td>
<td>0.273</td>
<td>0.031</td>
<td>21.031</td>
</tr>
<tr>
<td></td>
<td>(0.5)</td>
<td>(1.4)</td>
<td>(0.5)</td>
<td>(0.3)</td>
<td>(0.9)</td>
</tr>
<tr>
<td>d**</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1 d**</td>
<td>0.000</td>
<td>0.007***</td>
<td>-0.002**</td>
<td>0.000</td>
<td>0.051</td>
</tr>
<tr>
<td></td>
<td>(0.6)</td>
<td>(5.7)</td>
<td>(-2.1)</td>
<td>(0.2)</td>
<td>(1.5)</td>
</tr>
<tr>
<td>L2 d**</td>
<td>0.000</td>
<td>0.003**</td>
<td>-0.001</td>
<td>0.000</td>
<td>0.042</td>
</tr>
<tr>
<td></td>
<td>(0.1)</td>
<td>(2.2)</td>
<td>(-0.7)</td>
<td>(-0.1)</td>
<td>(1.1)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.088***</td>
<td>1.522***</td>
<td>-0.256*</td>
<td>-0.043**</td>
<td>3.287</td>
</tr>
<tr>
<td></td>
<td>(15.4)</td>
<td>(3.8)</td>
<td>(-1.7)</td>
<td>(-2.1)</td>
<td>(0.5)</td>
</tr>
<tr>
<td>Observations</td>
<td>1494</td>
<td>1494</td>
<td>1494</td>
<td>1494</td>
<td>1494</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.051</td>
<td>0.092</td>
<td>0.055</td>
<td>0.250</td>
<td>0.766</td>
</tr>
</tbody>
</table>

This table shows the coefficients of the VAR system estimated from Eq. (4). The system considers five endogenous variables reported in columns (1) to (5) respectively: the net dollar option order flow exposure to the volatility component (x*), to the underlying component (x*), and net SPY ETF order flow (x**); and the difference of the volatility process VIX (d*) and the price of the underlying measured by the SPY ETF (d*). The ordering of the columns corresponds to the ordering of the equations in the VAR model, and the coefficients set to missing identify the structural shocks. The letter L in the independent variable names represent the lag-operator. Inference is based on Newey–West standard errors with two lags that are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

on the underlying of $0.104 and $-0.131 dollars for the call and put, respectively. In basis points, these values are only 0.53 and -0.67.

We next convert these impacts on the underlying and volatility to obtain the price impact in the option. Column (10) shows the price impact of the individual option trades – thus ignoring cross-option price impact of the straddle – and these are 29.9 bps for the call, 42.1 bps for the put, yielding a value weighted portfolio average of 36 bps. As the price impact is linear, the impact for the average trade size is about 20 times smaller. The main result is column (12), which shows that the portfolio cost of 36 bps reduces the value of the underlying of $0.104 and $-0.131 dollars for the call and put, respectively.

We next convert these impacts on the underlying and volatility to obtain the price impact in the option. Column (10) shows the price impact of the individual option trades – thus ignoring cross-option price impact of the straddle – and these are 29.9 bps for the call, 42.1 bps for the put, yielding a value weighted portfolio average of 36 bps. As the price impact is linear, the impact for the average trade size is about 20 times smaller. The main result is column (12), which shows that the portfolio cost of 36 bps reduces the value of the underlying of $0.104 and $-0.131 dollars for the call and put, respectively.

Summarizing, the first result is that even the massive trades we consider in individual ATM options have a price impact of at most 36 bps, which is small relative to the average ATM option bid–ask spreads of 152 bps (Table 2, column (5)). Second, when looking at option portfolio trades, this price impact may shrink significantly depending on the extent that exposures to the underlying and volatility.

24 The put option has a larger price impact on the underlying because of the leverage effect: its vega exposure increases volatility, which in turn reduces the underlying. The reverse holds for the call option, where the positive impact of delta flow on the underlying gets partially reversed through the options increase on volatility.
expect that volatility is not affected by order flow exposures to the underlying. We now investigate whether these results can be

4.3. The common component in the underlying and the volatility process

Lastly, our finding that ETF order flow has a 3.4 times higher per dollar price impact than delta flow suggests a degree of market segmentation between SPX option markets and the SPY ETF. Option trades are much larger in terms of dollar volume and risk, which suggests that investors who prefer to trade in size use options. In contrast, the ETF market attracts relatively small but informed trades.

4.3. The common component in the underlying and the volatility process

The main analysis shows that delta and SPY order flow affect the volatility process, which is surprising because one would expect that volatility is not affected by order flow exposures to the underlying. We now investigate whether these results can be

Table 7
Hypothetical option portfolio price impact costs.

<table>
<thead>
<tr>
<th>(1) Price</th>
<th>(2) Strike</th>
<th>(3) Dollar cost ($1,000)</th>
<th>(4) Delta</th>
<th>(5) Vega</th>
<th>(6) $x^2$ ($100mln)</th>
<th>(7) $x^3$ ($1mln)</th>
<th>(8) dU (cents)</th>
<th>(9) dV (pp)</th>
<th>(10) Trade PI (bps)</th>
<th>(11) Portfolio PI (bps)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Straddle:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long ATM Call</td>
<td>16.88</td>
<td>1955</td>
<td>1688</td>
<td>0.63</td>
<td>141</td>
<td>1.24</td>
<td>1.24</td>
<td>10.4</td>
<td>–0.011</td>
<td>29.9</td>
</tr>
<tr>
<td>Long ATM Put</td>
<td>16.9</td>
<td>1955</td>
<td>1690</td>
<td>–0.38</td>
<td>140</td>
<td>–0.74</td>
<td>1.23</td>
<td>–13.1</td>
<td>0.015</td>
<td>42.1</td>
</tr>
<tr>
<td>Portfolio</td>
<td>3378</td>
<td></td>
<td>0.50</td>
<td>2.47</td>
<td>–2.7</td>
<td>0.004</td>
<td>36.0</td>
<td>1.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strangle</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long OTM put</td>
<td>7.7</td>
<td>1920</td>
<td>770</td>
<td>–0.16</td>
<td>103</td>
<td>–0.31</td>
<td>0.91</td>
<td>–6.9</td>
<td>0.005</td>
<td>21.1</td>
</tr>
<tr>
<td>Long OTM call</td>
<td>2.72</td>
<td>1990</td>
<td>272</td>
<td>0.18</td>
<td>76</td>
<td>0.36</td>
<td>0.67</td>
<td>1.9</td>
<td>–0.004</td>
<td>1.3</td>
</tr>
<tr>
<td>Portfolio</td>
<td>1042</td>
<td></td>
<td>0.04</td>
<td>1.58</td>
<td>–5.0</td>
<td>0.001</td>
<td>16.0</td>
<td>0.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Put (bear) spread</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long ITM Put</td>
<td>33.48</td>
<td>1990</td>
<td>3,348</td>
<td>–0.79</td>
<td>82</td>
<td>–1.54</td>
<td>0.72</td>
<td>–20.8</td>
<td>0.021</td>
<td>54.0</td>
</tr>
<tr>
<td>Short OTM put</td>
<td>7.7</td>
<td>1920</td>
<td>–770</td>
<td>–0.16</td>
<td>103</td>
<td>0.31</td>
<td>–0.91</td>
<td>6.9</td>
<td>–0.005</td>
<td>21.1</td>
</tr>
<tr>
<td>Portfolio</td>
<td>2578</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Call (bull) spread</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long ITM call</td>
<td>42.25</td>
<td>1920</td>
<td>4,225</td>
<td>0.84</td>
<td>104</td>
<td>1.64</td>
<td>0.92</td>
<td>16.3</td>
<td>–0.021</td>
<td>27.2</td>
</tr>
<tr>
<td>Short OTM call</td>
<td>2.72</td>
<td>1990</td>
<td>–272</td>
<td>0.18</td>
<td>76</td>
<td>–0.36</td>
<td>–0.67</td>
<td>–1.9</td>
<td>0.004</td>
<td>1.3</td>
</tr>
<tr>
<td>Portfolio</td>
<td>3953</td>
<td></td>
<td>1.28</td>
<td>0.25</td>
<td>14.4</td>
<td>–0.017</td>
<td>29.1</td>
<td>22.6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This table provides the expected price impact cost of the four most commonly used option portfolio strategies based on the VAR results in Table 5. Specifically, we consider a straddle, a strange and a put and a call spread. For each strategy, we measure the long-run cumulative price impact of the individual option trades and the portfolio trades, where the price impacts are netted or amplified. Specifically, we transform the option trades to delta and vega flows (columns (6) and (7)), which affect the underlying volatility process (columns (8) and (9)), and in turn the prices of the options (columns (10) and (11)). While most variables are self-explanatory, the variable Portfolio PI measures the price impact of the portfolio trade on the individual options, which incorporates the cross-option price impact of all option trades in the portfolio. As a benchmark, Trade PI measures the price impact of the single option trade, ignoring cross-option price impact. We calculate these results for option characteristics of ITM, ATM, and OTM call and put options of June 27, 2014 (the middle of the sample period), and take the latest trade before 12:00 pm. The options mature in 22 trading days (one month). The quantity of each option trade is a 1000 contracts, which each have a multiplier of 100. At the time, the value of the underlying was $1,954.2 and the volatility 0.088 (or 8.8%).

The estimated price impacts of delta and vega flow reflect this updating (or learning) process. Under this interpretation, we have estimated the price impact of volatility speculation, which is novel: a standard deviation shock to vega order flow ($16 million) increases the volatility by 0.14 standard deviations (0.07 percentage points). This price impact parameter can also be interpreted as the illiquidity of volatility speculation. The low volatility price impact we find means that volatility-related information is highly valuable because it can be exploited without moving the price much.

Learning and informed trading. To the extent that the estimated price impacts are permanent, the results are consistent with standard theories of informed trading (e.g., Kyle (1985), Glosten and Milgrom (1985)), with the extension that some investors are endowed with private information on the underlying or its volatility. They trade options to speculate on both information signals. While the market does not have this private information, it is aware of the general presence of informed traders, and rationally uses the observed option flows to update beliefs about the fundamental value and its volatility.

The estimated price impacts of delta and vega flow reflect this updating (or learning) process. Under this interpretation, we have estimated the price impact of volatility speculation, which is novel: a standard deviation shock to vega order flow ($16 million) increases the volatility by 0.14 standard deviations (0.07 percentage points). This price impact parameter can also be interpreted as the illiquidity of volatility speculation. The low volatility price impact we find means that volatility-related information is highly valuable because it can be exploited without moving the price much.

Price pressures. The price impacts may also reflect slow-moving transitory price pressures, that is, the analysis will capture temporary price changes that need more than three hours to die out. Under this interpretation, the results are consistent with inventory models of risk-averse market makers (see e.g., Ho and Stoll (1981)). For example, a market maker may hedge a trade in a given option by trading other options or the underlying, and accordingly create price pressures in those assets. If we assume that option prices depend only on changes in the underlying and volatility, then each market maker will decompose an option inventory position into a delta and vega exposure and calculate its risk position in the exact same fashion as in our model. In reality, option prices depend on other factors as well, including some noise, but our model should offer a good description of market maker inventory risks.

Market segmentation. Lastly, our finding that ETF order flow has a 3.4 times higher per dollar price impact than delta flow suggests a degree of market segmentation between SPX option markets and the SPY ETF. Option trades are much larger in terms of dollar volume and risk, which suggests that investors who prefer to trade in size use options. In contrast, the ETF market attracts relatively small but informed trades.

4.3. The common component in the underlying and the volatility process

The main analysis shows that delta and SPY order flow affect the volatility process, which is surprising because one would expect that volatility is not affected by order flow exposures to the underlying. We now investigate whether these results can be
explained by the leverage effect, i.e., the strong negative contemporaneous correlation between the underlying and the volatility process (Black, 1976). The Heston model specifically recognizes this correlation, but does not argue in which way the causality goes: Is it a shock from volatility to the underlying, or vice versa? While the causality does not matter for the pricing of an option, it is important to know where information originates when analyzing informed trading.

In this subsection, we take a simple approach to account for the leverage effect. We decompose the time series $d^u_t$ and $d^v_t$ into two components: the variation explained by the other variable and a residual, which we identify through the regressions:

$$d^u_t = c_1 + \sum_{l=0}^{\infty} \beta_1 l d^v_{t-l} + \varepsilon_{1,t}.$$  

$$d^v_t = c_2 + \sum_{l=0}^{\infty} \beta_2 l d^u_{t-l} + \varepsilon_{2,t}.$$  

These equations are simultaneously determined. The predicted values of each captures the variation that is common in both variables, which we call $d^{c,u}_t$ and $d^{c,v}_t$. We interpret these as the variation in $d^u_t$ or $d^v_t$ explained by the leverage effect.\textsuperscript{25} The residuals, then, capture the variation that is not caused by the leverage effect, which we call $d^{r,u}_t$ and $d^{r,v}_t$.

Accordingly, we estimate the VAR model of Eq. (4) twice, where we replace $d^u_t$ and $d^v_t$ (i) with $d^{c,u}_t$ and $d^{c,v}_t$ to use the common variation in both variables; and (ii) with $d^{r,u}_t$ and $d^{r,v}_t$ to use the residual variation. All else remains the same.

Table 8 shows the OLS results analogous to Table 6 (the IRFs are provided in Figs. 2 and 3). For brevity, we only present the regressions with the underlying and volatility as dependent variables (the other equations in the VAR remain nearly identical). Columns (1) and (2) show the VAR specification using the common variations, $d^{c,u}_t$ and $d^{c,v}_t$. We see that $x^\Delta$, $x^{\Delta \text{SPY}}$, and $x^{\text{SPY}}$ all significantly predict $d^{c,v}_t$. In column (2), $x^\Delta$ and $x^{\text{SPY}}$ no longer significantly predict $d^{c,u}_t$, because all their explanatory power has already been captured through the contemporaneous regressor $d^{c,v}_t$ in that equation. To see this, we repeat regression (2) but omit

\textsuperscript{25} We cannot isolate the leverage effect from the two series without stronger assumptions. In fact, the predicted values $d^{c,u}_t$ and $d^{c,v}_t$ are both noisy estimates of the leverage effect, where the former is affected by $\varepsilon_{2,t}$ and the latter by $\varepsilon_{1,t}$ in Eqs. (6) and (7). To see this, a shock to $\varepsilon_{1,t}$ increases $d^v_t$ and therefore $d^{c,v}_t$, but it does not affect $d^{c,u}_t$. 

---

**Fig. 1. Impulse response functions VAR: the impact of order flows on the underlying and volatility returns** In this figure, we plot the IRFs of order flow shocks to the difference of the underlying ($d^u$) and volatility ($d^v$). The $d^u$ is based on the SPY ETF quoted midpoint (in cents) and the $d^v$ is based on the VIX extracted from the cross section of option prices (in percentage points). The net order flow (buyer initiated minus seller initiated) in each option is decomposed into dollar order flow exposure to the underlying ($\Delta x^\Delta$) based on its delta, and volatility ($\Delta x^\text{vega}$) based on its vega. The order flows are aggregated across options with all strike prices, expiration dates, and puts and calls. The order flow in the SPY ETF is also added in a separate equation. The IRFs are based on the VAR system of Eq. (4) and corresponds to the results in Table 6.
A. Kaeck et al.

Fig. 2. Impulse response functions VAR: common variation in $d^u$ and $d^v$. In this figure, we plot the IRFs of the VAR model. It is similar to Fig. 1, except that we replace $d^u$ and $d^v$ by the predicted value of the regressions:

$$d^u_t = c_1 + \sum_{l=0}^{L} \beta_{1,l} d^v_{t-l} + \epsilon_{1,t},$$

$$d^v_t = c_2 + \sum_{l=0}^{L} \beta_{2,l} d^u_{t-l} + \epsilon_{2,t}.$$  
The predicted values capture the variation in $d^u$ ($d^v$) explained by $d^v$ ($d^u$), and thus the variation that reflects the leverage effect. This specification corresponds to columns (1) and (2) of Table 5.

Fig. 3. Impulse response functions VAR: residual variation in $d^u$ and $d^v$. In this figure, we plot the IRFs of the VAR model. It is similar to Fig. 1, except that we replace $d^u$ and $d^v$ by the residuals of the regressions:

$$d^u_t = c_1 + \sum_{l=0}^{L} \beta_{1,l} d^v_{t-l} + \epsilon_{1,t},$$

$$d^v_t = c_2 + \sum_{l=0}^{L} \beta_{2,l} d^u_{t-l} + \epsilon_{2,t}.$$  
The residuals capture the variation in $d^u$ ($d^v$) not explained by $d^v$ ($d^u$), and thus filter out the leverage effect. This specification corresponds to columns (4) and (5) of Table 5.
We decompose $d^c$ and $d^v$ into two components: the variation explained by each other and a residual, identified by running:

\[ d^c_i = c_1 + \sum_{t=1}^{T} \beta_{1c} d^v_{i,t} + \epsilon_{1c}, \]
\[ d^v_i = c_2 + \sum_{t=1}^{T} \beta_{2v} d^c_{i,t} + \epsilon_{2v}. \]

We denote the predicted value of the first equation by $d^{c*}$, as it represents the common variation explained by $d^v$; and denote the residual $\epsilon_{1c}$ by $d^{c**}$. The predicted value from the second equation is called $d^{v*}$ and the residual $d^{v**}$. The OLS regressions below correspond to the five-equation VAR model of Table 6, but based on the common variation (columns (1) and (2)) or the residual variation (columns (4) and (5)). Columns (3) and (6) are similar to columns (2) and (5), respectively, but do not have the contemporaneous $d^v$ term.

We observe that the effect of vega flow on residual volatility is only transitory, and becomes insignificant after three hours. This result is consistent with inventory models, and not with models of informed trading where the price impact is permanent. This test does lack statistical power because the volatility process is noisily observed due to large option spreads and tick sizes. This noise is mostly captured by the residual component in volatility, since the common component has little noise to the extent that the underlying has little noise.
multiplied by the factor (change in underlying price or change in volatility). This analysis uses data of all individual options (one

\( \Delta d^e \) and \( \nu d^e \) are the signed mid-quote changes in \( \Delta n, t \) and \( \nu_n, t \), respectively. The options midpoint price and the change in the underlying and the volatility are sampled at the hourly frequency (the data set is balanced). We add two lags to be consistent with the previous analyses. The sample uses the full sample of option data of 2014, including options with all expiration dates, strike prices and puts and calls. Column (1) shows results for the full sample, and columns (2)-(10) for various subsets of options sorted by ttm and moneyness. We estimate the equations in the VAR

\[
\Delta n, t = \beta_0 + \sum_{l=1}^{2} \beta_{l1} \Delta_{n,t-l} d^e_{-l} + \sum_{l=0}^{2} \beta_{l2} \nu_{n,t-l} d^e_{-l} + \epsilon_{n,t},
\]

(8)

where \( d^e_n \) is the change in the mid-quote of option \( n \), and the terms \( \Delta_n d^e_n \) and \( \nu_n d^e_n \) represent the options exposure (delta or vega) multiplied by the factor (change in underlying price or change in volatility). This analysis uses data of all individual options (one

26 We have nine subgroups, but by adding the aggregate delta and vega flows we loose one subgroup, which becomes the base case. The test is not affected by which group becomes the base case, because we test whether the cumulative coefficients of all groups are identical. With two lags in the system, we jointly test whether the eight sums (of three coefficients of the delta flows or vega flows) are equal to zero. We perform the test for the \( d^e \) and \( d^e \) equations in the VAR.
The VAR used one observation per day-hour. For consistency with the VAR model, we include two lags in the regression. If the structure imposed by a two-factor option pricing model is correct, the regression should yield a) an R-Squared of one, and b) coefficients \( \beta_1 = 0, \sum_{l=1}^2 \beta_{1,l} = 1, \) and \( \sum_{l=2}^2 \beta_{2,l} = 1. \) Indeed, in this theoretical case the empirical specification would perfectly explain the actual changes in option prices.

Table 9 provides our estimation results. Column (1) shows the full sample regression results. We obtain an R-squared of 99.2%, which suggests that our model does an excellent job at summarizing the information content in option price changes. We also find that \( \sum_{l=1}^2 \beta_{1,l} = 0.997 \) and \( \sum_{l=2}^2 \beta_{2,l} = 0.899, \) which both are close to one from an economical point of view. This also confirms that option returns are strongly affected by changes in volatility, and that options are useful assets to speculate on changes in volatility. The \( t \)-statistic on coefficient \( \delta \times d^\delta \) is extremely large (5551) due to the high R-Squared and the sheer size of the data (851,051 observations). For this reason, we do reject equality to one for both \( \sum_{l=1}^2 \beta_{1,l} \) and \( \sum_{l=2}^2 \beta_{2,l} \) (the \( t \)-statistics are 14.6 and 37.3, respectively). This is easily explained by a small model misspecification. Column (1) further shows that the lagged coefficients are smaller, but still statistically significant. This motivates the use of the lags in the system. Compared to Bakshi et al. (2000), the two-factor model works much better with more recent data because markets have become more efficient. Using SPX option data from 1994, they find coefficients of \( \beta_1 = 0.80, \beta_2 = 0.41, \) and an R-squared of 59%.

We repeat the exercise for the nine subsets of options sorted by ttm and moneyness. In general, the model works very well. We see it performs slightly weaker for options with a low moneyness and short ttm (column (2)). In this case, the \( \sum_{l=1}^2 \beta_{1,l} = 0.612 \) and \( \sum_{l=2}^2 \beta_{2,l} = 1.35. \) These coefficients are likely affected by the model misspecification of the Heston model, which, for example, does not allow for discontinuous jump moves in the underlying price equation. In particular, it has been shown that for pricing short ttm options modeling a jump component yields better pricing performance compared to a pure stochastic volatility model (see Eraker (2004)). Further, this subset contains options with very low prices, and we know that the microstructure noise is more severe here as tick sizes are relatively larger. The model works better for the remaining columns, which all have an R-squared exceeding 92%.

From an economical standpoint, the imposed structure fits the data well. However, the rejection of equality to one for \( \sum_{l=1}^2 \beta_{1,l} \) and \( \sum_{l=2}^2 \beta_{2,l} \) means that actual option price changes differ from what our model predicts. This implies that either the deltas and vegas contain errors; or options price changes have transitory components; or that the true data generating process for option prices contains additional factors (see, e.g., Christoffersen et al. (2009) or Bardgett et al. (2019)). As a consequence, there is some bias in the delta and vega order flows used in the VAR model. This issue could be tackled by using a more advanced option pricing model, but this is beyond the scope of this paper. However, we see no economic channel how any misspecification of the two-factor structure would alter the findings in the VAR. In fact, any misspecification would most likely bias the results against finding price impacts.

5. Higher frequencies

We have shown that at the one-hour frequency, delta and vega order flows predict changes in the underlying and volatility. A limitation of the identification in the VAR model is the restriction that one endogenous variable can affect another contemporaneously, but that the latter cannot affect the former. This assumption seems tenuous at the hourly frequency, and in this section we investigate whether we can extend the analysis to the half-hour and one-minute frequency.

Estimating the VAR model for higher frequencies such as one-minute intervals poses a non-trivial challenge since a much larger number of lags needs to be added to the VAR equations to cover a comparable time horizon as for the one-hour frequency. A large number of lags implies a large number of parameters, which in turn makes estimation procedures unstable. To deal with this issue, we follow the procedure proposed in Hashbrouck (2019), which uses polynomial distributed lag (PDL) functions to decrease the number of model parameters. In this approach, the time horizon covered by the many lags is divided into a few subintervals, for each of which a PDL function is fitted. This procedure reduces the number of parameters significantly by restricting the individual lag coefficients to lie on the PDL function. After estimating the PDL functions we can easily calculate the parameters of the individual lags and impulse response functions.

Our main finding is that the VAR analysis provides weaker results at the half-hour frequency, and insignificant results at the one-minute frequency. Specifically, the coefficients and IRFs of delta flow and vega flow on changes in the underlying and volatility shrink, and they turn insignificant at the one-minute frequency (see Figs. 4 and 5, respectively). After careful investigation, we conclude that at higher frequencies microstructure noise becomes too severe, weakening all effects of interest. Microstructure noise is a classical form of measurement error and biases the OLS coefficients [Buccheri et al. (2019) discuss this issue for price discovery regressions]. Options have large tick sizes and bid–ask spreads, which prevents prices to adjust smoothly from adjusting changes in the underlying asset or the volatility. Only at sufficiently low frequencies are the changes in the two factors large enough (relative to bid–ask spreads) to induce option price changes.

We conduct two analyses to reveal the effect of microstructure noise on the estimates. First, Table 10 shows that the bid–ask spreads and tick sizes of options are large compared to the standard deviation of option price changes at high frequency. The tick size ranges from $0.05 to $0.10 (depending on the option price level).\(^27\) The bid–ask spreads however, calculated as the median across option categorized into eight subsets by option price level, range from $0.35 for options priced under a dollar to $1.56 for options priced between $20 and $40. This means the spreads are typically between 7 and 15 ticks. The spread values are similar to the standard deviation of price changes at the hourly level, which range from $0.14 to $1.45 for the same subsets of options. At the one-minute frequency, however, the standard deviations of price changes range between $0.05 and $0.17, which is about 7 to 10 times smaller than the bid–ask spread. This implies that at the one-minute level, the changes in the two price factors are so small that the implied option price changes often fall within the bid and ask quotes and do not update frequently.

\(^{27}\) Currently, SPX options have a tick size of $0.05 when the price is below $3.00, and $0.10 otherwise. Source: http://www.cboe.com.
Fig. 4. Impulse response functions VAR: half-hour frequency This figure is identical to Fig. 1, but shows results for data sampled at the half-hour frequency. For comparison to the main specification (at the hourly level), we add four lags in the VAR system.

Fig. 5. Impulse response functions VAR: one-minute frequency This figure is identical to Fig. 1, but shows results for data sampled at the one-minute frequency. For comparison to the main specification (at the hourly level), we add ten lags in the VAR system and iterate the IRFs for 20 steps (minutes).
and delta flow) will be attributed to the latter, leaving less explanatory power for vega flow. This effect can be seen by comparing

Given that vega flow is ordered last in the VAR, any contemporaneous correlation between vega flow and SPY flow (or vega flow

problematic: the noise makes it more difficult to distinguish between the impact of order flows on the underlying and the volatility.

these option prices, which occurs more slowly. An additional issue is that the microstructure noise makes the leverage effect more

not update frequently, they will not quickly reflect order flow information. In turn, the volatility process itself is extracted from

Measurement error caused by microstructure noise directly biases coefficients towards zero. Further, as quoted option prices do

structure appears to break down at the highest frequencies.

(5) and (8), where the R-squared is only 42.1%, 28.8%, and 19.1%, respectively. The two-factor structure is in essence a data

reduction technique that summarizes the information in many option prices and order flows by a handful of components. This

Table 10
Option tick sizes versus standard deviation of price changes.

<table>
<thead>
<tr>
<th>Option price category</th>
<th>0–1</th>
<th>1–5</th>
<th>5–10</th>
<th>10–20</th>
<th>20–40</th>
<th>40–100</th>
<th>100–500</th>
<th>500–∞</th>
<th>Full sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median Tick Size ($)</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>Median Spread ($)</td>
<td>0.35</td>
<td>0.70</td>
<td>1.27</td>
<td>1.56</td>
<td>2.30</td>
<td>2.94</td>
<td>2.77</td>
<td>2.31</td>
<td></td>
</tr>
<tr>
<td>Number of Options</td>
<td>257</td>
<td>951</td>
<td>397</td>
<td>432</td>
<td>433</td>
<td>676</td>
<td>1823</td>
<td>935</td>
<td>5854</td>
</tr>
</tbody>
</table>
| Average standard deviation of dollar price changes at different sampling frequencies
| Hour                  | 0.14| 0.19| 0.46  | 0.83  | 1.43  | 2.65   | 4.66    | 7.75  | 3.23        |
| Half-hour             | 0.10| 0.13| 0.32  | 0.59  | 1.02  | 1.9    | 3.45    | 5.75  | 2.37        |
| Minute                | 0.05| 0.03| 0.06  | 0.1   | 0.17  | 0.29   | 0.44    | 0.49  | 0.2         |

For each option, we calculate the tick size, the average dollar bid–ask spread, and the standard deviation of the price changes (in dollars) sampled at the one-hour, half-hour, and one-minute frequency. We then divide options into eight price categories, and we report the median of each variable across the options in each category. This differs from Table 2, where spreads are averaged over all trades.

Table 11
Explanatory power of price and volatility of the S&P 500 on option price changes at the one-minute frequency.

<table>
<thead>
<tr>
<th></th>
<th>Full sample</th>
<th>Subsets</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avg TTM (years)</td>
<td>0.1769</td>
<td></td>
</tr>
<tr>
<td>Avg Moneyness (S/K)</td>
<td>1.1156</td>
<td></td>
</tr>
<tr>
<td>(\Delta')</td>
<td>0.632***</td>
<td>0.245***</td>
</tr>
<tr>
<td>(L1) (\Delta')</td>
<td>0.910</td>
<td>(9.5)</td>
</tr>
<tr>
<td>(L2) (\Delta')</td>
<td>0.247***</td>
<td>0.081***</td>
</tr>
<tr>
<td>(L3) (\Delta')</td>
<td>0.031***</td>
<td>0.015</td>
</tr>
<tr>
<td>(\nu')</td>
<td>0.187***</td>
<td>0.043***</td>
</tr>
<tr>
<td>(L1) (\nu')</td>
<td>0.025***</td>
<td>0.014***</td>
</tr>
<tr>
<td>(L2) (\nu')</td>
<td>0.007*</td>
<td>0.014</td>
</tr>
<tr>
<td>(L3) (\nu')</td>
<td>0.004</td>
<td>0.026</td>
</tr>
<tr>
<td>Observations</td>
<td>10,169,421</td>
<td>393,753</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.479</td>
<td>0.421</td>
</tr>
</tbody>
</table>

This table is identical to Table 9, but uses data sampled at the one-minute frequency. We regress option price changes on the particular option's delta times the changes in the underlying (\(\Delta'\)) and the options vega times the change in the volatility (\(\nu'\)). We add ten lags of these variables. Due to the sheer size, we only use data of the first two months of 2014 rather than the whole year (results are unaffected). Column (1) shows results for the full sample, and columns (2)-(10) for various subsets of options sorted by time-to-maturity (ttm) and moneyness. The first rows show the average ttm and moneyness of the particular sample. Inference is based on Newey–West standard errors with two lags. The superscripts ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

The direct consequence of the microstructure noise is that the two-factor structure we impose has weak explanatory power at the highest frequencies. We repeat the analysis of Section 4.4 at the one-minute frequency, and find in Table 11 that the two factors explain no more than 47.9% of the variation in the full sample. Moreover, the coefficients on the terms \(\Delta'\) and \(\nu'\) (summed over the lags) lie much further away from 1. These issues are worse for the subsets of the options with low moneyness (columns (2), (5) and (8)), where the R-squared is only 42.1%, 28.8%, and 19.1%, respectively. The two-factor structure is in essence a data reduction technique that summarizes the information in many option prices and order flows by a handful of components. This structure appears to break down at the highest frequencies.

The poor performance of the two-factor model at the highest frequency also implies weaker results of the VAR model. Measurement error caused by microstructure noise directly biases coefficients towards zero. Further, as quoted option prices do not update frequently, they will not quickly reflect order flow information. In turn, the volatility process itself is extracted from these option prices, which occurs more slowly. An additional issue is that the microstructure noise makes the leverage effect more problematic: the noise makes it more difficult to distinguish between the impact of order flows on the underlying and the volatility. Given that vega flow is ordered last in the VAR, any contemporaneous correlation between vega flow and SPY flow (or vega flow and delta flow) will be attributed to the latter, leaving less explanatory power for vega flow. This effect can be seen by comparing...
the minute-frequency analysis in Fig. 5, where, compared to the hourly analysis in Fig. 1, the $x^{\text{PP}}$ gets a stronger predictive power at the expense of $x^i$ and $x^d$. Overall, we find that the hourly frequency we use is the best compromise between a high-frequency analysis and avoiding the adverse effects of market microstructure noise in option data.

6. Robustness analyses

We performed a number of robustness analyses, which we summarize in this section. Detailed results are in the Online Appendix.

**Ordering of the equations.** An important but somewhat arbitrary choice is the ordering of the equations in the structural VAR. This holds particularly at the hourly frequency, where essentially all series occur simultaneously. The ordering determines the direction of the contemporaneous correlation between two variables: only one is allowed to affect the other, but not the reverse. Without this structure, the system is not identified. The ordering in turn determines the structural innovations, the regression coefficients, and the impulse response functions. While five equations allow for 120 possible combinations, we note that for the price impact results, the order of $d^u$ and $d^v$ does not matter since they both appear after the order imbalance variables. There are two alternative orderings that deserve further attention. First, we estimate a version of our model where we put $x^{\text{PP}}$ before $x^d$ and $x^i$ (see Table 2, Panel B, in the Internet Appendix). In this case, the information contemporaneously captured in $x^{\text{PP}}$ and $x^i$ is attributed to the latter. Accordingly, the coefficient of $x^i$ on $d^u$ is estimated more precisely (a higher $t$-stat). The long-run effects however are similar to the main specification, both in terms of economic magnitudes and significance levels. Second, in Panel C we consider an alternative ordering $x^{\text{PP}}, x^i, x^d$, that switches delta and vega flow compared to Panel B. This change appears to have a negligible impact on the estimated coefficients.

**The forward price.** In the setup of our main analysis, we use the ETF SPY price to proxy for the value of the underlying asset. An alternative is to extract the forward price from the option data, which is noisier but not confounded by dividend yields paid before option expiration.\(^{28}\) We re-estimate the VAR using the forward price as the underlying instead and find nearly identical results (see Figure A3, in the Internet Appendix). This is reassuring, yet we prefer the main specification as the ETF price typically leads in price discovery (Hasbrouck, 2003). At the one-hour frequency, however, this channel is of no concern.

**Price changes versus returns.** We estimate the VAR with the return on the underlying and volatility, instead of price differences. After appropriately scaling the coefficients, the results are virtually identical to those in the main specification (see Figure A4, in the Internet Appendix). We chose to report the version with price differences, as it corresponds more naturally to the analysis in Section 4.4.

**The role of the SPY ETF order flow.** We estimate a four equation version of the VAR model, where we drop the SPY ETF order flow equation from the model. In this case, the results are generally stronger in terms of statistical significance (although the point estimates of the coefficients are similar). Not only do we estimate fewer parameters, also note that the ETF order flow is correlated with option order flows. When adding ETF flow in the main specification of the model, it absorbs part of the option flow’s predictive power (see Figure A5, in the Internet Appendix).

**Limitations:** The analysis is based on data of SPX options and the SPY ETF. In reality, traders can trade other investment vehicles to obtain exposure to the S&P 500 (through the S&P 500 future for example) or its volatility (through options on the VIX or options on the SPY for example). We do not have these data and thus are unable to investigate to which extent investors use such sources to trade on their information. Nonetheless, the main contribution of the paper is the model, which is flexible and allows for easy integration of other order flow sources. One approach is to calculate the delta and vega order flows in these additional assets and sum them with the current variables. An alternative is to incorporate separate equations for these sources, which allows for tests of differences of coefficients to determine which order flows have higher price impacts.

7. Conclusion

We offer a novel framework to estimate the price impact of the aggregate option market. To our knowledge, we are the first to propose a single model that captures trading in potentially hundreds of options, with differing strike prices, expiration dates, and types (put or call). We impose economic structure on the data using theoretical option pricing models, which takes into account the high cross-option correlations in returns, order flows, and liquidity. We are also the first to disentangle the leverage effect when studying informed trading in the underlying and the volatility.

Our main result is that SPX option price impacts on the underlying and the volatility are small, especially compared to effective spreads. For a typical at-the-money call trade, the price impact is less than 4 bps, compared to an effective spread of 169 bps. Further, if we analyze portfolio trades, rather than individual option trades, price impacts can easily shrink tenfold depending on the extent that the delta and vega order flow exposures cancel out.

This raises an important question: Why are option spreads so large compared to the very small price impact? Traditional explanations suggest that spreads depend on asymmetric information, market maker inventory effects, fixed order processing costs, and market maker rents. But the first two frictions also cause a price impact, which according to our results seem small. We believe it is unlikely that the latter two frictions are large for the highly competitive and liquid SPY ETF market. As such, the large observed spreads become even more puzzling.

---

\(^{28}\) We back out the underlying price process from option prices as follows. For each $t$ on a one-minute grid, we collect every put/call pair with identical strike price and identical maturity date and use market mid quotes to solve put–call parity for the unobservable futures prices. At a given time $t$ and for a fixed option maturity $T$, we average the implied futures prices over all available strikes for which a put and a call price is available to obtain the option implied futures price $F_{t,T}$. We use linearly-interpolated rates from the OptionMetrics zerocurve file as a proxy for the risk-free rate of interest. We then average the changes over all short-term futures in our sample to calculate $d^u$. 

---
Appendix. Heston model details

Model definition. The Heston (1993) model has become the most important benchmark in the option pricing literature. Its main theoretical advantage stems from the mathematical tractability of its characteristic function (see Duffie et al. (2000)). Under the model assumptions, the S&P 500 index $S$ and its volatility $\nu$ are described by the following stochastic differential equations under the risk-neutral measure $\mathbb{Q}$:

\[
\begin{align*}
    dS_t &= (r - q)S_t dt + S_t \nu_t dW_t^1 \\
    d\nu_t &= \kappa (\Theta - \nu_t) dt + \sigma \nu_t \left( \rho dW_t^1 + \sqrt{1 - \rho^2} dW_t^2 \right),
\end{align*}
\]

where $W^1$ and $W^2$ are standard Brownian motions under $\mathbb{Q}$. Structural parameters are given by $\kappa > 0$, $\Theta > 0$, $\nu > 0$, and $\rho \in [-1, 1]$. The risk-free rate is denoted $r$ and $q$ is the continuous dividend yield.

Due to the affine model structure, European call and put prices can be calculated by standard Fourier inversion methods (see below) and we denote $P(t, S_t, \nu_t, K, \tau, \omega)$ as the time-$t$ price of a put or call option ($\omega = 1$ for a call, $\omega = -1$ for a put) with strike $K$ and maturity $\tau > t$. We collect all structural parameters and the risk-free rate in the vector $\Theta$ but drop the dependence of option prices on the parameter vector for notational simplicity. It follows from Ito’s lemma that:

\[
\begin{align*}
    dP(t, S_t, \nu_t, K, \tau, \omega) &= \frac{\partial P(t, S_t, \nu_t, K, \tau, \omega)}{\partial S_t} dS_t + \frac{\partial P(t, S_t, \nu_t, K, \tau, \omega)}{\partial \nu_t} d\nu_t + dt \text{-terms.}
\end{align*}
\]

Due to the analytical tractability of the option pricing function $P$, the first partial derivatives with respect to $S$ and $\nu$ can also be calculated analytically and we denote these as:

\[
\begin{align*}
    \Delta(t, S_t, \nu_t, K, \tau, \omega) &\equiv \frac{\partial P(t, S_t, \nu_t, K, \tau, \omega)}{\partial S_t} \\
    \nu(t, S_t, \nu_t, K, \tau, \omega) &\equiv \frac{\partial P(t, S_t, \nu_t, K, \tau, \omega)}{\partial \nu_t}.
\end{align*}
\]

To relate changes in option prices to observable variables, we exploit the theoretical relation between the spot volatility and the VIX index. It is straightforward to show that under our model assumptions, the squared VIX index is a linear function of spot and its volatility $\nu$:

\[
\begin{align*}
    \nu(t, S_t, \nu_t, K, \tau, \omega) = \frac{1}{2} \log \left( \frac{S_t}{S_0} \right) - \frac{1}{2} \log \left( \frac{\nu_t}{\nu_0} \right).
\end{align*}
\]

For the Heston model, the logarithm of the characteristic function under the risk-neutral measure $\mathbb{Q}$ is given by:

\[
\log \Psi^\mathbb{Q}(u, \tau, S_t, \nu_t) \equiv \log E^\mathbb{Q} \left[ e^{iu \log S_{\tau}} | F_t \right] = A(u, \tau) + B(u, \tau) \nu_t^2 + u \log S_t,
\]

where $A$ and $B$ are complex-valued functions, $i = \sqrt{-1}$ and $E^\mathbb{Q} \left[ \cdot | F_t \right]$ denotes the risk-neutral $F_t$-conditional expectation. For expositional clarity, the dependence of all functions on the parameter set of the model $\Theta$ is suppressed. Following Bates (2006), the price of a European call option is given by:

\[
\begin{align*}
    P(t, S_t, \nu_t, K, \tau, \omega = 1) &= e^{-\omega \tau} F_{t, \tau} - e^{-\omega (T - \tau) K} \left( \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \Re \left[ \frac{e^{-iu \log K \Psi^\mathbb{Q}(u, T - t, S_t, \nu_t)}}{iu(1 - iu)} \right] du \right),
\end{align*}
\]

where $F_{t, \tau} = e^{(r-q)(T-t)S_t}$ and $\Re$ denote the real part of a complex number. Partial derivatives of the call price are given by:

\[
\begin{align*}
    \frac{\partial P(t, S_t, \nu_t, K, \tau, \omega = 1)}{\partial S_t} &= e^{-\omega (T - t) \nu} - e^{-\omega (T - t) K} \frac{1}{\pi} \int_0^\infty \Re \left[ \frac{iu e^{-iu \log K \Psi^\mathbb{Q}(u, T - t, S_t, \nu_t)}}{iu(1 - iu)} \right] du,
\end{align*}
\]

and

\[
\begin{align*}
    \frac{\partial P(t, S_t, \nu_t, K, \tau, \omega = 1)}{\partial \nu_t} &= -\frac{2}{\pi} \nu \times e^{-\omega (T - t) K} \frac{1}{\pi} \int_0^\infty \Re \left[ \frac{B(u, T - t) e^{-iu \log K \Psi^\mathbb{Q}(u, T - t, S_t, \nu_t)}}{iu(1 - iu)} \right] du.
\end{align*}
\]

The theoretical value of the VIX index with maturity $\tau_v = 30/365$ can be recovered from the characteristic function. One can show that for the Heston model:

\[
VIX_{\tau} = \sqrt{\frac{\Theta}{\kappa} \left( 1 - \frac{1 - e^{\kappa \tau_v}}{\tau_v \kappa^2} \right) + \frac{1 - e^{\kappa \tau_v}}{\tau_v \kappa^2} \nu^2_v}. \tag{A.6}
\]

Empirical estimation. In order to compute the delta and vega for all option transactions in our data set, we need to calibrate the Heston model to market data. There is no standard methodology in the literature regarding the estimation of model parameters of option pricing models. We adopt a simple calibration procedure and estimate model parameters for every trading day in our sample using intradaily option quotes. To this end, we minimize the mean root square error of all available OTM options with maturities from 7 to 180 days as follows:

\[
RMSE_t = \arg \min_{\Theta} \sqrt{\frac{1}{N_t} \sum_{i=1}^{N_t} \left( P^{\text{market}}(t_i, K_i, \tau_i, \omega_i) - P(t_i, S_{t_i}, \nu_{t_i}, K_i, \tau_i, \omega_i, \Theta) \right)^2}, \tag{A.7}
\]
where $N_i$ is the number of available option quotes on day $t_i$, $t_i$ are set to minutely intervals from 9:45 am to 15:45 pm (on day $t$) and $P^{ma}(t_i;K_i,T_{i},o_i)$ denotes the market mid-quote of option $i$. We use the shortest option maturity to invert the theoretical VIX formula above to obtain $v_i$. To limit the number of option contracts in our calibration, we restrict the calibration to the most liquid contracts, OTM contracts, and contracts with short to medium maturity. For the calculation of deltas and vegas, we use calibrated parameters from the previous trading day.

Our calibration procedure has two distinct features. First, our methodology circumvents the problem of filtering the latent variance process, as we calibrate the model to both option prices and the VIX index simultaneously. Alternative approaches such as those in Broadie et al. (2007) or Christoffersen et al. (2010) rely on filtering techniques or the calibration of $v_i$, which leads to additional complexity in the calibration algorithm. And second, our recalibration allows us to be robust to changing market dynamics. While the Heston model may be rejected because of the imposed structure (see for instance Broadie et al. (2007) or Christoffersen et al. (2010)), our recalibration ensures that empirical results are not particularly sensitive to possible model misspecification. To calculate deltas and vegas for option transactions, we use parameters estimated using option data on the previous day, so our procedure is out-of-sample. While parameters in the calibration may change over time, our assumption is that the Heston model provides a reasonable way of separating price from volatility risk using model calibrations from the previous trading day. Following the empirical set-up in Bakshi et al. (2000), we provide strong empirical evidence in Table 9 that this approach covers more than 98% of the variation in option prices in our sample. Hence, Eq. (A.3) in combination with our Heston model implementation provides a highly accurate description of option market prices, and offers the main benefit of a reduction of variables to only price and volatility risk.

Appendix B. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.finmar.2021.100675.

References


